RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/896,301

DATE: 12/21/2001
TIME: 14:29:46

Input Set : N:\Crf3\RULE60\09896301.txt
Output Set: N:\CRF3\12212001\I896301.raw

<110> APPLICANT: Cosgrove, Daniel J
McQueen-Mason, Simon
Guiltinan, Mark J
Shcherban, Tatyana
Shi, Jun
<120> TITLE OF INVENTION: PURIFIED EXPANSIN PROTEINS
<130> FILE REFERENCE: 1194/1C114US3
<140> CURRENT APPLICATION NUMBER: 09/896,301
<141> CURRENT FILING DATE: 2001-06-29
<150> PRIOR APPLICATION NUMBER: 09/092,160
<151> PRIOR FILING DATE: 1998-06-05
<150> PRIOR APPLICATION NUMBER: 08/440,517
<151> PRIOR FILING DATE: 1995-05-12
<150> PRIOR APPLICATION NUMBER: 08/242,090
<151> PRIOR FILING DATE: 1994-05-12
<150> PRIOR APPLICATION NUMBER: 08/060,944
<151> PRIOR FILING DATE: 1993-05-12
<160> NUMBER OF SEQ ID NOS: 7
<170> SOFTWARE: PatentIn Ver. 2.1
<210> SEQ ID NO: 1
<211> LENGTH: 681
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: cDNA cucumber expansin
<400> SEQUENCE: 1

```
gactacggtg gctggcagag cggccacgcc acctttatg gtggtggtga cgcatctggc 60
accatgggtg gagcttgtgg gtatgggaat ttatacagcc aagggtatgg cacgaacacg 120
gtggcgctga gcactgcgct atttaacaat ggattaagtt gtggtgcttg cttcgaaatg 180
acttgtacaa acgaccctaa atggtgcctt ccgggaacta ttagggtcac tgccaccaac 240
ttttgccctc ctaactttgc tctccctaac aacaatggtg gatggtgcaa ccctcctctc 300
caacacttcg acatggctga gcctgccttc cttcaaatcg ctcaataccg agctggtatc 360
gtccccgtct cctttcgtag ggtaccatgt atgaagaaag gtggagtgag gtttacaatc 420
aatggccact catacttcaa cctcgttttg atcacaaacg tcggtggcgc aggcgacgtc 480
cactctgtgt cgataaaggg gtctcgaact ggatggcaat ccatgtctag aaattggggc 540
caaaactggc aaagcaacaa ctatctcaat ggccaaggcc tttcctttca agtcactctt 600
agtgatggtc gcactctcac tgcctataat ctcgttcctt ccaattggca atttggccaa 660
acctatgaag gccctcaatt c                                          681
```

<210> SEQ ID NO: 2
<211> LENGTH: 228
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: rice expansin
<220> FEATURE:
<221> NAME/KEY: UNSURE

RAW SEQUENCE LISTING DATE: 12/21/2001
PATENT APPLICATION: US/09/896,301 TIME: 14:29:46

Input Set : N:\Crf3\RULE60\09896301.txt
Output Set: N:\CRF3\12212001\I896301.raw

```
127 <222> LOCATION: 211
129 <223> OTHER INFORMATION: Xaa is unknown or other.
133 <400> SEQUENCE: 2
135 Ala Gly Gly Gly Trp Val Asn Ala His Ala Thr Phe Tyr Gly Gly Gly
137   1               5                  10                  15
141 Asp Ala Ser Gly Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn Leu Tyr
143              20                  25                  30
147 Ser Gln Gly Tyr Gly Thr Asn Thr Ala Ala Leu Ser Thr Ala Leu Phe
149          35                  40                  45
153 Asn Asn Gly Leu Ser Cys Gly Ala Cys Phe Glu Ile Arg Cys Gln Asn
155      50                  55                  60
159 Asp Gly Lys Trp Cys Leu Pro Gly Ser Ile Val Val Thr Ala Thr Asn
161  65                  70                  75                  80
165 Phe Cys Pro Pro Asn Asn Ala Leu Pro Asn Asn Ala Gly Gly Trp Cys
167                  85                  90                  95
171 Asn Pro Pro Gln Gln His Phe Asp Leu Ser Gln Pro Val Phe Gln Arg
173              100                 105                 110
177 Ile Ala Gln Tyr Arg Ala Gly Ile Val Pro Val Ala Tyr Arg Arg Val
179          115                 120                 125
183 Pro Cys Val Arg Arg Gly Gly Ile Arg Phe Thr Ile Asn Gly His Ser
185      130                 135                 140
189 Tyr Phe Asn Leu Val Leu Ile Thr Asn Val Gly Gly Ala Gly Asp Val
191 145                 150                 155                 160
195 His Ser Ala Met Val Lys Gly Ser Arg Thr Gly Trp Gln Ala Met Ser
197                 165                 170                 175
201 Arg Asn Trp Gly Gln Asn Trp Gln Ser Asn Ser Tyr Leu Asn Gly Gln
203              180                 185                 190
207 Ser Leu Ser Phe Lys Val Thr Thr Ser Asp Gly Gln Thr Ile Val Ser
209          195                 200                 205
W--> 213 Asn Asn Xaa Ala Asn Ala Gly Trp Ser Phe Gly Gln Thr Phe Thr Gly
215      210                 215                 220
219 Ala His Val Arg
221 225
227 <210> SEQ ID NO: 3
229 <211> LENGTH: 222
231 <212> TYPE: PRT
233 <213> ORGANISM: Artificial Sequence
237 <220> FEATURE:
239 <223> OTHER INFORMATION: Description of Artificial Sequence: rice expansin
243 <220> FEATURE:
245 <221> NAME/KEY: UNSURE
247 <222> LOCATION: (14)..(58)
249 <223> OTHER INFORMATION: Xaa is unknown or other.
253 <400> SEQUENCE: 3
W--> 255 His Met Gly Pro Trp Ile Asn Ala His Ala Thr Phe Tyr Xaa Xaa Gly
257   1               5                  10                  15
W--> 261 Asp Ala Xaa Xaa Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn Leu Tyr
263              20                  25                  30
267 Ser Gln Gly Tyr Gly Leu Glu Thr Ala Ala Leu Ser Thr Ala Leu Phe
```

RAW SEQUENCE LISTING DATE: 12/21/2001
PATENT APPLICATION: US/09/896,301 TIME: 14:29:46

Input Set : N:\Crf3\RULE60\09896301.txt
Output Set: N:\CRF3\12212001\I896301.raw

```
     269          35                  40                  45
W--> 273 Asp Gln Gly Leu Ser Cys Gly Ala Cys Xaa Glu Leu Met Cys Val Asn
     275      50                  55                  60
     279 Asp Pro Gln Trp Cys Ile Lys Gly Arg Ser Ile Val Val Thr Ala Thr
     281  65                  70                  75                  80
     285 Asn Phe Cys Pro Pro Gly Gly Ala Cys Asp Pro Pro Asn His His Phe
     287                  85                  90                  95
     291 Asp Leu Ser Gln Pro Ile Tyr Glu Lys Ile Ala Leu Tyr Lys Ser Gly
     293             100                 105                 110
     297 Ile Ile Pro Val Met Tyr Arg Arg Val Arg Cys Lys Arg Ser Gly Gly
     299         115                 120                 125
     303 Ile Arg Phe Thr Ile Asn Gly His Ser Tyr Phe Asn Leu Val Leu Val
     305     130                 135                 140
     309 Thr Asn Val Gly Gly Ala Gly Asp Val His Ser Val Ser Met Lys Gly
     311 145                 150                 155                 160
     315 Ser Arg Thr Lys Trp Gln Leu Met Ser Arg Asn Trp Gly Gln Asn Trp
     317                 165                 170                 175
     321 Gln Ser Asn Ser Tyr Leu Asn Gly Gln Ser Leu Ser Phe Val Val Thr
     323             180                 185                 190
     327 Thr Ser Asp Arg Arg Ser Val Val Ser Phe Asn Val Ala Pro Pro Thr
     329         195                 200                 205
     333 Trp Ser Phe Gly Gln Thr Tyr Thr Gly Gly Gln Phe Arg Tyr
     335     210                 215                 220
     341 <210> SEQ ID NO: 4
     343 <211> LENGTH: 227
     345 <212> TYPE: PRT
     347 <213> ORGANISM: Artificial Sequence
     351 <220> FEATURE:
     353 <223> OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
     355       expansin
     359 <220> FEATURE:
     361 <221> NAME/KEY: UNSURE
     363 <222> LOCATION: (2)..(227)
     365 <223> OTHER INFORMATION: Xaa is unknown or other.
     369 <400> SEQUENCE: 4
W--> 371 Lys Xaa Ser Val Ala Gln Ser Ala Phe Ala Thr Phe Tyr Gly Gly Lys
     373   1               5                  10                  15
     377 Asp Gly Ser Cys Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn Leu Tyr
     379              20                  25                  30
     383 Asn Ala Gly Tyr Gly Leu Tyr Asn Ala Ala Leu Ser Ser Ala Leu Phe
     385          35                  40                  45
     389 Asn Asp Gly Ala Met Cys Gly Ala Cys Tyr Thr Ile Thr Cys Asp Thr
     391      50                  55                  60
     395 Ser Gln Thr Lys Trp Cys Lys Pro Gly Gly Asn Ser Ile Thr Ile Thr
     397  65                  70                  75                  80
W--> 401 Ala Thr Asn Leu Cys Xaa Pro Asn Trp Ala Leu Pro Ser Asn Ser Gly
     403                  85                  90                  95
W--> 407 Gly Trp Cys Asn Pro Pro Leu Xaa His Phe Asp Met Ser Gln Pro Ala
     409             100                 105                 110
```

RAW SEQUENCE LISTING DATE: 12/21/2001
PATENT APPLICATION: US/09/896,301 TIME: 14:29:46

Input Set : N:\Crf3\RULE60\09896301.txt
Output Set: N:\CRF3\12212001\I896301.raw

```
     413 Trp Glu Asn Ile Ala Val Tyr Gln Ala Gly Ile Val Pro Val Asn Tyr
     415         115                 120                 125
W--> 419 Lys Arg Val Pro Xaa Gln Arg Ser Gly Gly Ile Arg Phe Ala Ile Ser
     421     130                 135                 140
     425 Gly His Asp Tyr Phe Glu Leu Val Thr Val Thr Asn Val Gly Gly Ser
     427 145                 150                 155                 160
     431 Gly Val Val Ala Gln Met Ser Ile Lys Gly Ser Asn Thr Gly Trp Met
     433                 165                 170                 175
     437 Ala Met Ser Arg Asn Trp Gly Ala Asn Trp Gln Ser Asn Ala Tyr Leu
     439             180                 185                 190
     443 Ala Gly Gln Ser Leu Ser Phe Ile Val Gln Leu Asp Asp Gly Arg Lys
     445         195                 200                 205
W--> 449 Val Thr Ala Trp Asn Xaa Ala Pro Xaa Asn Trp Leu Xaa Xaa Xaa Xaa
     451     210                 215                 220
W--> 455 Xaa Xaa Xaa
     457 225
     463 <210> SEQ ID NO: 5
     465 <211> LENGTH: 225
     467 <212> TYPE: PRT
     469 <213> ORGANISM: Artificial Sequence
     473 <220> FEATURE:
     475 <223> OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
     477       expansin
     481 <400> SEQUENCE: 5
     483 Asp Asn Gly Gly Trp Glu Arg Gly His Ala Thr Phe Tyr Gly Gly Ala
     485   1               5                  10                  15
     489 Asp Ala Ser Gly Thr Met Gly Gly Ala Cys Gly Tyr Gly Asn Leu His
     491             20                  25                  30
     495 Ser Gln Gly Tyr Gly Leu Gln Thr Ala Ala Leu Ser Thr Ala Leu Phe
     497         35                  40                  45
     501 Asn Ser Gly Gln Lys Cys Gly Ala Cys Phe Glu Leu Thr Cys Glu Asp
     503     50                  55                  60
     507 Asp Pro Glu Trp Cys Ile Pro Gly Ser Ile Ile Val Arg Tyr Asn Leu
     509 65                  70                  75                  80
     513 Ala Asn Phe Ala Leu Ala Asn Asp Asn Gly Gly Trp Cys Asn Pro Pro
     515                 85                  90                  95
     519 Leu Lys His Phe Asp Leu Ala Glu Pro Ala Phe Leu Gln Ile Ala Gln
     521             100                 105                 110
     525 Tyr Arg Ala Gly Ile Val Pro Val Ala Phe Arg Arg Val Pro Cys Glu
     527         115                 120                 125
     531 Lys Gly Gly Gly Ile Arg Phe Thr Ile Asn Gly Asn Pro Tyr Phe Asp
     533     130                 135                 140
     537 Leu Val Leu Ile Thr Asn Val Gly Gly Ala Gly Asp Ile Arg Ala Val
     539 145                 150                 155                 160
     543 Ser Leu Lys Gly Ser Lys Thr Asp Gln Trp Gln Ser Met Ser Arg Asn
     545                 165                 170                 175
     549 Trp Gly Gln Asn Trp Gln Ser Asn Thr Tyr Leu Arg Gly Gln Ser Leu
     551             180                 185                 190
     555 Ser Phe Gln Val Thr Asp Ser Asp Gly Arg Thr Val Val Ser Tyr Asp
```

RAW SEQUENCE LISTING DATE: 12/21/2001
PATENT APPLICATION: US/09/896,301 TIME: 14:29:46

Input Set : N:\Crf3\RULE60\09896301.txt
Output Set: N:\CRF3\12212001\I896301.raw

```
557         195                 200                 205
561 Val Val Pro His Asp Trp Gln Phe Gly Gln Thr Phe Glu Gly Gly Gln
563     210                 215                 220
567 Phe
569 225
575 <210> SEQ ID NO: 6
577 <211> LENGTH: 226
579 <212> TYPE: PRT
581 <213> ORGANISM: Artificial Sequence
585 <220> FEATURE:
587 <223> OTHER INFORMATION: Description of Artificial Sequence: Arabidopsis
589       expansin
593 <400> SEQUENCE: 6
595 Asp Tyr Ser Ser Trp Gln Ser Ala His Ala Thr Phe Tyr Gly Gly Gly
597   1               5                  10                  15
601 Asp Ala Ser Gly Thr Met Gly Gly Thr Cys Gly Tyr Gly Asn Leu Tyr
603             20                  25                  30
607 Ser Thr Gly Tyr Thr Asn Thr Ala Ala Leu Ser Thr Val Leu Phe Asn
609          35                  40                  45
613 Asp Gly Ala Ala Cys Arg Ser Cys Tyr Glu Leu Arg Cys Asp Asn Asp
615      50                  55                  60
619 Gly Gln Trp Cys Leu Pro Gly Ser Val Thr Val Thr Ala Thr Asn Leu
621  65                  70                  75                  80
625 Cys Pro Pro Asn Tyr Ala Leu Pro Asn Asp Asp Gly Gly Trp Cys Asn
627                  85                  90                  95
631 Pro Pro Arg Pro His Phe Asp Met Ala Glu Pro Ala Phe Leu Gln Ile
633             100                 105                 110
637 Gly Val Tyr Arg Ala Gly Ile Val Pro Val Ser Tyr Arg Arg Val Pro
639         115                 120                 125
643 Cys Val Lys Lys Gly Gly Ile Arg Phe Thr Ile Asn Gly His Ser Tyr
645     130                 135                 140
649 Phe Asn Leu Val Leu Val Thr Asn Val Ala Gly Pro Gly Asp Val Gln
651 145                 150                 155                 160
655 Ser Val Ser Ile Lys Gly Ser Ser Thr Gly Trp Gln Pro Met Ser Arg
657                 165                 170                 175
661 Asn Trp Gly Gln Asn Trp Gln Ser Asn Ser Tyr Leu Asp Gly Gln Ser
663             180                 185                 190
667 Leu Ser Phe Gln Val Ala Val Ser Asp Gly Arg Thr Val Thr Ser Asn
669         195                 200                 205
673 Asn Val Val Pro Ala Gly Trp Gln Phe Gly Gln Thr Phe Glu Gly Gly
675     210                 215                 220
679 Gln Phe
681 225
687 <210> SEQ ID NO: 7
689 <211> LENGTH: 227
691 <212> TYPE: PRT
693 <213> ORGANISM: Artificial Sequence
697 <220> FEATURE:
699 <223> OTHER INFORMATION: Description of Artificial Sequence: cucumber
```

VERIFICATION SUMMARY DATE: 12/21/2001
PATENT APPLICATION: US/09/896,301 TIME: 14:29:47

Input Set : N:\Crf3\RULE60\09896301.txt
Output Set: N:\CRF3\12212001\I896301.raw

L:213 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:255 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:261 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:273 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:371 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:401 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:407 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:419 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:449 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:455 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4